To: Beck, Nancy[Beck.Nancy@epa.gov]; Yamada, Richard (Yujiro)[yamada.richard@epa.gov]; Rodan, Bruce[rodan.bruce@epa.gov]; Bahadori, Tina[Bahadori.Tina@epa.gov]; Dourson, Michael[dourson.michael@epa.gov]; Morris, Jeff[Morris.Jeff@epa.gov]; Henry, Tala[Henry.Tala@epa.gov]; Orme-Zavaleta, Jennifer[Orme-Zavaleta.Jennifer@epa.gov]; Hanley, Mary[Hanley.Mary@epa.gov]; Bertrand, Charlotte[Bertrand.Charlotte@epa.gov]; Scheifele, Hans[Scheifele.Hans@epa.gov]; Lavoie, Emma[Lavoie.Emma@epa.gov]; Camacho, Iris[Camacho.Iris@epa.gov]; Bateson, Thomas[Bateson.Thomas@epa.gov]; Kraft, Andrew[Kraft.Andrew@epa.gov]
From: Thayer, Kris
Sent: Wed 12/20/2017 8:36:10 AM
Subject: Background information on study scoring in systematic review

Thanks again for the meeting today. Hopefully both of our programs will benefit from the robust discussion.

As noted today, I am headed out of the country so will unlikely be able to make another meeting this week. However, I wanted to quickly share some of the systematic review guidance materials that we referenced as not supportive of the type of scoring approach we understand is being developed for TSCA.

2011 Institute of Medicine report (where the TSCA definition of systematic review is taken from) https://www.nap.edu/read/13059/chapter/5#132

p. 132 In recent years, systematic review teams have moved away from scoring systems to assess the quality of individual studies towards a focus on the components of quality and risk of bias (Juni, 1999). Quality scoring systems have not been validated. Studies assessed as excellent quality using one scoring method may be subsequently assessed as lower quality using another scoring method (Moher et al., 1996). Moreover, with an emphasis on risk of bias, the SR more appropriately assesses the quality of study design and conduct rather than the quality of reporting.

Cochrane Handbook (the most recognized source of systematic review guidance)
http://handbook-5-1.cochrane.org/chapter_8/8_3_3_quality_scales_and_cochrane_reviews.htm

8.3.3 Quality scales and Cochrane reviews

The use of scales for assessing quality or risk of bias is explicitly discouraged in Cochrane reviews. While the approach offers appealing simplicity, it is not supported by empirical evidence (Emerson 1990, Schulz 1995b). Calculating a summary score inevitably involves

assigning 'weights' to different items in the scale, and it is difficult to justify the weights assigned. Furthermore, scales have been shown to be unreliable assessments of validity (Jüni 1999) and they are less likely to be transparent to users of the review. It is preferable to use simple approaches for assessing validity that can be fully reported (i.e. how each trial was rated on each criterion).

The underlying limitations to scoring are considered applicable to a wide variety of evidence (clinical trials, observational epidemiological studies, animal bioassays, etc.) and whether the assessment is narrative or a meta-analysis. The NCEA comments on the TSCA approach collected and shared by Emma make the same points as outlined in the IOM and Cochrane documents. Thus, we have a convergence of opinions among practitioners of systematic review within NCEA and best practice recommendations from the other fields. In contrast, I am not aware of any current systematic review guidance that recommends the use of numerical scoring. At best, scoring approaches are tolerated when other tools do not exist to evaluate the evidence (e.g., in vitro evidence). However, this is not the case for epidemiological and experimental animal evidence.

If we could resolve the scoring issue, then it would be MUCH easier to have approaches that are consistent.

As an aside, the approach to evaluating epidemiology studies within IRIS is based on the methods outlines in this publication "***ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions***", which has been cited almost 175 times since being published in BMJ in 2016. A couple of the IRIS epidemiologists are collaborating with this group to develop a version of this tool targeted to assessing environmental and occupational exposures. This is one strategy we are taking to ensure that the SR methods implemented in IRIS are in sync with the best practices in the broader SR community.

Happy holidays! -k

Kristina Thayer, Ph.D.

Director, Integrated Risk Information System (IRIS) Division

National Center for Environmental Assessment, NCEA

ORD, USEPA

Mail Code: B243-01

Building: Bldg B (Room B211I)

Research Triangle Park, NC 27711

(919) 541-0152 RTP

(202) 564-1771 Ronald Reagan Building (room 51203)

Skype: kristina.thayer

thayer.kris@epa.gov